The extragalactic submillimetre population: predictions for the SCUBA Half-Degree Extragalactic Survey (SHADES)

Article (Published Version)

Mon. Not. R. Astron. Soc. **359,** 469–480 (2005) doi:10.1111/j.1365-2966.2005.08899.x

The extragalactic submillimetre population: predictions for the SCUBA Half-Degree Extragalactic Survey (SHADES)

Eelco van Kampen,[1,2⋆] Will J. Percival,[1] Miller Crawford,[1] James S. Dunlop,[1] Susie E. Scott,[1] Neil Bevis,[3] Seb Oliver,[3] Frazer Pearce,[4] Scott T. Kay,[3] Enrique Gaztañaga,[5,6] David H. Hughes[5] and Itziar Aretxaga[5]

[1]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*
[2]*Institute for Astrophysics, University of Innsbruck, Technikerstr. 25, A-6020 Innsbruck, Austria*
[3]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH*
[4]*School of Physics & Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD*
[5]*Instituto Nacional de Astrofísica, Optica y Electrónica, Apt. Postal 51 y 216, Puebla, Pue, Mexico*
[6]*Institut d'Estudis Espacials de Catalunya, Edifici Nexus, Gran Capita 2-4, desp. 201, 08034 Barcelona, Spain*

Accepted 2005 February 4. Received 2005 February 1; in original form 2004 July 7

ABSTRACT

We present predictions for the angular correlation function and redshift distribution for SHADES, the SCUBA Half-Degree Extragalactic Survey, which will yield a sample of around 300 submillimetre sources in the 850-μm waveband in two separate fields. Complete and unbiased photometric redshift information on these submillimetre sources will be derived by combining the SCUBA data with (i) deep radio imaging already obtained with the Very Large Array, (ii) guaranteed-time *Spitzer* data at mid-infrared wavelengths, and (iii) far-infrared maps to be produced by BLAST, the Balloon-borne Large-Aperture Submillimeter Telescope. Predictions for the redshift distribution and clustering properties of the final anticipated SHADES sample have been computed for a wide variety of models, each constrained to fit the observed number counts. As we are dealing with around 150 sources per field, we use the sky-averaged angular correlation function to produce a more robust fit of a power-law shape $w(\theta) = (\theta/A)^{-\delta}$ to the model data. Comparing the predicted distributions of redshift and of the clustering amplitude A and slope δ, we find that models can be constrained from the combined SHADES data with the expected photometric redshift information.

Key words: surveys – galaxies: evolution.

1 INTRODUCTION

The understanding of galaxy formation and evolution is making rapid progress because of a combination of a wealth of new observational data at a wide range of wavelengths and redshifts, and an increased understanding of which physical processes underlie the formation and evolution of galaxies. However, an important problem with most current galaxy formation models is that it is difficult to establish whether a set of model parameters that produces a good match to observations is unique. The main reason for this is that, for most models, there are degeneracies amongst the various free parameters. As a significant fraction of observational data used to constrain the model parameters are obtained from our local Universe, the 'uniqueness problem' can be resolved by comparing model predictions and observations at high redshift, which in many respects is independent of a comparison at low redshift.

One major difference between low and high redshifts is the frequency and intensity of major mergers: at high redshifts they occur far more often and are more intense as the participating galaxies are likely to be more gas-rich than their low-redshift counterparts. They are also expected to be dust-enshrouded at the peak of their burst of star formation, which means that they are most easily detected in the submillimetre or far-infrared wavebands.

For this purpose, highly valuable observational data will be provided by the SCUBA Half-Degree Extragalactic Survey [SHADES: see http://www.roe.ac.uk/ifa/shades and Mortier et al. (2005) for details]. This survey, which commenced in 2002 December, has been designed to cover 0.5 deg^2 to a 3.5σ detection limit of $S_{850\,\mu m}$ = 8 mJy, split between two 0.25-deg^2 fields. The two survey areas, the Lockman Hole East and the Subaru–*XMM* Deep Field (SXDF), have been selected on the basis of low galactic confusion at submillimetre wavelengths, and the wealth of existing or

⋆E-mail: eelco.v.kampen@uibk.ac.at

anticipated supporting multi-frequency data from radio to X-ray wavelengths.

In addition the SCUBA data will be combined with data from the Very Large Array (VLA), the *Spitzer Space Telescope* and the Balloon-borne Large-Aperture Submillimeter Telescope (BLAST) [see http://chile1.physics.upenn.edu/blastpublic and Devlin (2001) for details], which will undertake a series of nested extragalactic surveys at 250, 350 and 500 μm. This experiment will significantly extend the wavelength range, sensitivity and area of existing ground-based extragalactic submillimetre surveys (Hughes et al. 2002).

It is anticipated that spectroscopic redshifts will ultimately be obtained for a substantial fraction of the SHADES sources (e.g. Chapman et al. 2003, 2005). However, the key point for the work presented here is that, even where optical/near-infrared spectroscopy is impossible, the long-wavelength data provided by the combined SCUBA + VLA + *Spitzer* + BLAST data set will yield photometric redshifts for *all* sources with uncertainties of $\delta z < 0.5$ (Aretxaga, Hughes & Dunlop 2005). This offers a unique powerful way of providing the complete and unbiased redshift and spectral energy distribution (SED) information required to measure the clustering properties of submillimetre sources, and the cosmic history of dust-enshrouded star formation that takes place in very massive starbursts with inferred star formation rates of the order of 1000 $M_\odot\,yr^{-1}$ (Scott et al. 2002).

These massive starbursts could be associated with the formation of the progenitors of massive ellipticals if sustained for a significant amount of time (up to 1 Gyr). However, SCUBA sources could also be associated with bright, but short-lived, bursts of intense star formation occurring in more modest galaxies drawn from the high-redshift galaxy population already discovered at optical/ultraviolet wavelengths (Adelberger & Steidel 2000, and many others). If the bright SCUBA sources are indeed the progenitors of massive ellipticals then they are likely to be more strongly clustered than when drawn from the population of less massive galaxies. This is an inevitable result of gravitational collapse from Gaussian initial density fluctuations: the rare high-mass peaks are strongly biased with respect to the mass (Kaiser 1984).

There is abundant evidence that this bias does occur at high redshift: the correlations of Lyman-break galaxies at $z \simeq 3$ are almost identical to those of present-day field galaxies, even though the mass must be much more uniform at early times. Moreover, the correlations increase with ultraviolet luminosity (Giavalisco & Dickinson 2001), reaching scalelengths of $r_0 \simeq 7.5\,h^{-1}$ Mpc, 1.5 times the present-day value. Daddi et al. (2000) find a trend of clustering with colour for extremely red objects (EROs), reaching $r_0 \simeq 11\,h^{-1}$ Mpc for $R - K > 5$, which corresponds to fluctuations in projected number density that are ∼unity on the scale of the SCUBA field of view, falling to 10 per cent rms on 1° scales. For SHADES to detect the clustering properties of bright submillimetre sources over comoving scales reaching $\geqslant 10$ Mpc, the survey needs to cover a significant fraction of a square degree. At the time of writing, the survey is set to reach half a square degree within three years, and is making good progress towards achieving that goal.

The direct predecessor of SHADES was the 8-mJy survey of Scott et al. (2002; see also Ivison et al. 2002). The correlation function for SCUBA sources derived from this survey alone did not yield a significant detection of clustering, even though the large uncertainties meant that it was still consistent with the strong clustering displayed by EROs. There are nevertheless good reasons for believing the SCUBA source population to be highly clustered, and some observational evidence for this is now found (Blain et al. 2004). In particular, cross-correlation with X-ray sources (Almaini et al. 2003) and Lyman-break galaxies (Webb et al. 2003) yields clearly significant detections of clustering.

Scott et al. (in preparation) have recently performed a combined clustering analysis on the three main existing blank-field SCUBA surveys [the 8-mJy survey, the CUDSS survey by Webb et al. (2003) and the Hawaii survey by Barger, Cowie & Sanders (1999)] to determine whether the existing data are capable of revealing significant clustering within the submillimetre population alone. Even though this analysis is based on combining data from several small fields, it has yielded the first significant (5σ) measurement of submillimetre source clustering on scales ≃0.5–2 arcmin, of a strength that does indeed appear comparable to that found by Daddi et al. (2000) for EROs. Interestingly, if the integral constraint (see Section 4.2 for its definition) is varied as a free parameter, the inferred clustering in fact becomes stronger than that displayed by the ERO population.

Finally, Blain et al. (2004) have found tentative evidence for a clustering length of $6.9 \pm 2.1\,h^{-1}$ Mpc (comoving) for those submillimetre galaxies for which they could obtain spectroscopic redshifts. As this was possible only for sources that are also detected at radio wavelengths, their sample is incomplete and likely to be biased. Furthermore, Adelberger (2005) argued that the method used is prone to systematic errors and is unnecessarily noisy.

The paper is organized as follows. Section 2 provides an overview of SHADES and its main aims. Section 3 describes the various models used to make predictions, which are compared with each other in Section 4. This section also presents the actual predictions for SHADES, and we discuss these in Section 5.

2 SHADES: A WIDE-AREA SUBMILLIMETRE SURVEY WITH REDSHIFT INFORMATION

The science goals of SHADES are to help to answer three fundamental questions about galaxy formation: What is the cosmic history of massive dust-enshrouded star formation activity? Are SCUBA sources the progenitors of present-day massive ellipticals? What fraction of SCUBA sources harbour a dust-obscured active galactic nucleus (AGN)? The aim of this paper is to review and compare the predictions of various existing models for the bright submillimetre population, and to consider how they can be best tested and constrained by the final SHADES data set, and thus help to answer the first two questions. The third question is not addressed in this paper, as it involves a detailed analysis of the combined radio, mid-infrared and X-ray properties of the SHADES sources.

An important property of SHADES is having meaningful redshift estimates, which provide vital information for estimating the bolometric luminosity of the sources, and hence the cosmic history of energy output from dust-enshrouded star formation activity. Redshift information also holds the key to measuring the clustering properties of the submillimetre source population. Although precise spectroscopic measurements of the redshift of a sample of SHADES sources will be possible if reliable radio/optical/infrared counterparts can be identified and readily followed up with 10-m-class optical telescopes (e.g. Chapman et al. 2003, 2005), in practice one will not be able to derive this information for the majority of sources in the survey. However, combination of the SHADES and BLAST data will allow the use of submillimetre photometric redshift techniques, yielding crude estimates ($\delta z < 0.5$) for individual sources detected in both surveys.

A Monte Carlo based photometric redshift technique has been designed by Hughes et al. (2002) and tested by Aretxaga et al. (2003, 2005). Here submillimetre photometric information is combined with prior information on the population, such as the number

counts and the likely evolution of the luminosity function of dust-enshrouded galaxies, to weight the output redshifts provided by a large sample of template SEDs. These SEDs represent the wide range of temperatures, dust emissivities and luminosities found in nearby infrared-bright galaxies.

Even though the redshift distributions are relatively wide, the detailed information on the shape of the distribution, combined with a large number of sources, provides a powerful statistical measurement of population properties such as the parent redshift distribution and the global star formation rate. While, naively, these measurements might seem insufficiently crude, the combination of the redshift distributions of hundreds of sources can indeed measure the history of star formation of the galaxies detected in more than two submillimetre bands ($L_{\rm FIR} > 2 \times 10^{13}\,{\rm L}_{\odot}$) with an accuracy of $\sim$20 per cent (Hughes et al. 2002).

While simulations show that photometric redshift estimates detected only from submillimetre data have errors of the order of 0.5 (Hughes et al. 2002), it has been shown empirically that the inclusion of additional photometric information provided by detections or upper limits at 1.4 GHz (from the VLA) and at 170–70 μm (from the *Spitzer Space Telescope*) increases the accuracy of the photometric redshifts to $\pm$ 0.3 (Aretxaga et al. 2005).

3 FOUR ALTERNATIVE MODELS OF THE EXTRAGALACTIC SUBMILLIMETRE POPULATION

Four different models for the clustering of SCUBA galaxies are presented: a 'simple merger' model, a 'hydrodynamical' model, a 'stable clustering' model and a 'phenomenological' model. Some of these are designed especially with SHADES in mind, while for other models the SCUBA predictions are just part of a range of predictions. The models also vary in the level of complexity, and in the underlying assumptions, including the choices for the cosmological parameters, even though differences in the latter are minor compared with the fundamental differences between the models.

The aim of this paper is not to perform a detailed comparison between these models, or between models and data, but simply to present predictions for a diverse range of realistic models. The goal is to study the ability of SHADES to measure clustering, and establish its capability to distinguish between models.

3.1 A simple merger model

The simple merger model is included in order to help to determine the important processes at work in the creation of SCUBA galaxies. The underlying premise of this model is that 8-mJy SCUBA galaxies are formed by obscured star formation driven by the violent merger between two galaxy-sized haloes. The emission is assumed to be above the 8-mJy detection threshold for a lifetime $t_{\rm life}$ after the galaxy haloes have merged. No direct link is made between the luminosity of the SCUBA galaxy and the properties of the merger except that a lower limit is placed on the final mass of haloes that contain a detectable 8-mJy SCUBA source. In other words, a Poisson sampling of massive halo mergers is assumed to form bright SCUBA galaxies. We have adopted a mass limit of $10^{13}\,{\rm M}_{\odot}$, corresponding to 'radio galaxy' mass haloes.

Halo mergers were found in a 256^3 N-body simulation run within a comoving $(100\,h^{-1}\,{\rm Mpc})^3$ box using GADGET, a publicly available parallel tree code (Springel, Yoshida & White 2001). Cosmological parameters were assumed to have their concordance values ($\Omega_{\rm m} = 0.3$, $\Omega_{\Lambda} = 0.7$, $h = 0.70$ and $n_{\rm s} = 1$), and the power spectrum normalization was set at $\sigma_8 = 0.9$. Outputs from the simulation were obtained at 434 epochs, separated approximately uniformly in time, and haloes were found at each epoch using a standard friends-of-friends routine with linking length $b = 0.2$. New haloes were defined to be haloes with >50 per cent of the constituent particles not having previously been recorded in a halo of equal or greater mass. Of these, the halo was said to have been created by a major merger if there were two progenitors at the previous time output that had mass between 25 and 75 per cent of the final mass.

Obtaining the right number density of SCUBA sources is limited by the definition of merging used, the lifetime of emission above the detection threshold, and the proportion of mergers that result in SCUBA sources. We therefore simply assume that all of the mergers, defined as above, result in a luminous SCUBA source, and allow $t_{\rm life}$ to vary to give $\sim$300 sources in 0.5 deg^2. Because the density of high-mass ($>10^{13}\,{\rm M}_{\odot}$) mergers is low, obtaining the correct number density of SCUBA sources requires a relatively long lifetime $t_{\rm life} = 8 \times 10^8$ yr.

Mock SCUBA catalogues for a 0.5-deg^2 survey are calculated by placing a (comoving) light cone through an array of simulation boxes. This is done by selecting output time-steps such that corresponding redshifts are separated by a box length in comoving coordinates. Boxes are reflected, rotated and translated randomly to reduce the artificial correlation between neighbouring boxes inherent in using a single simulation; this necessarily reduces real correlations due to structures that would cross boxes. Mergers that occur less than the model lifetime before the time corresponding to their luminosity distance are flagged as potential SCUBA sources, and their angular positions and redshifts are recorded in order to create mock catalogues.

Obviously, while this model does predict both the spatial distribution and redshift-space distribution of the SCUBA sources, it does not predict the luminosity function. In fact, we note that following successful comparison between analytic theory and numerical simulations, both the redshift-space distribution and the spatial distribution of SCUBA galaxies in this model could have been accurately estimated analytically (Percival, Miller & Peacock 2000; Percival et al. 2003).

3.2 A hydrodynamical model

At the heart of the model is a simulation from Muanwong et al. (2002) that is an adaptive particle–particle, particle–mesh code incorporating smoothed particle hydrodynamics (SPH). The underlying code is HYDRA (Couchman, Thomas & Pearce 1995) with the addition of a standard pair-wise artificial viscosity (Thacker et al. 2000). The cosmological model is $\Omega_{\rm m} = 0.35$, $\Omega_{\Lambda} = 0.65$, $h = 0.71$, $\sigma_8 = 0.9$, $\Omega_{\rm b} = 0.019\ h^{-2}$. The simulation used here employs a box of comoving size $(100\,h^{-1}\,{\rm Mpc})^3$ with 160^3 dark matter particles and 160^3 gas particles, and is evolved between $50 > z > 0$ in approximately 2000 time-steps. The simulations have various components: non-interacting dark matter; gas; 'star-like' (same as gas, but forming stars); and 'galaxy fragments', which are collisionless. Evolution of the various components is as follows: all particles evolve under gravity; gas can adiabatically heat and cool; gas can also radiatively cool; at $\rho/\bar{\rho} > 500$, $T < 12\,000$ K gas particles become 'star-like' (at this point all the mass is deemed to have been converted into stars). An aggregation of 13 or more close 'star-like' particles become a 'fragment'. Fragments may accrete more star-like particles but do not merge.

As a complete model of galaxy formation this simulation has a number of strengths and weaknesses. It provides a self-consistent treatment of large-scale structure and galaxy evolution. However, the limited resolution and the arbitrary solution to the 'cooling catastrophe' necessitated by this limit its validity. For the present purpose the full power of the simulation is not used; it serves as an ingredient to a more phenomenological model.

The first step is to construct a 'galaxy fragment' light-cone in the usual way (as described in Section 3.1). With the SHADES sample area it is not necessary to use more than a single box transverse to the line of sight.

The redshift distribution of the fragments in this cone $(\mathrm{d}N/\mathrm{d}z)_{\mathrm{frag}}$ is measured using redshift bins of uniform width $\Delta z = 0.7$; however, there are many more fragments in each bin than would be detectable. Each fragment is treated as the possible location of a SCUBA source and is selected based upon its star formation rate (SFR), which is measured as the mass per unit time of 'star-like' particles accreted to this fragment (averaged over the last output time-step). The required number of fragments with the greatest SFR are selected from each bin such that the redshift distribution matches a particular model: $(\mathrm{d}N/\mathrm{d}z)_{\mathrm{SCUBA}}$. For this paper, the analytical form of Baugh, Cole & Frenk (1996) is adopted with a median redshift of 2.3 and the normalization such as to give 300 sources in the full 0.5-deg^2 sample size. Hence the redshift distribution produced is not derived from the hydrodynamical simulations and this model merely makes a reasonable choice as to which fragments SHADES will include, and, for these, encodes the positional information from the simulations.

3.3 A stable clustering model

This is the model of Hughes & Gaztañaga (2000), in which a single output from a *N*-body simulation that fits well the local spatial correlation function as measured by APM (Gaztañaga & Baugh 1998) is used to generate a population of SCUBA galaxies in a light-cone. This corresponds to assuming stable clustering, i.e. a constant spatial correlation function in comoving space. Fixing the spatial correlation function does not imply that we also fix the angular correlation, as that depends on light-cone geometry, luminosity evolution, and the redshift selection function. The prescription for galaxy formation corresponds to the assumption that the probability for finding a galaxy somewhere within the light-cone is simply proportional to the local dark matter density, with the total number of galaxies normalized to the surface density required for a given flux limit. Although the redshift distribution is (exponentially) cut off beyond $z = 6$, this model contains the highest redshift SCUBA galaxies of all models considered in this paper.

This model has been used in the photometric redshift estimation technique of Aretxaga et al. (2003), to constrain sample size and depth given the correlation length, and to test correlation function measurements from surveys with relatively small sky coverage (Gaztañaga & Hughes 2001).

3.4 A phenomenological model

The phenomenological galaxy formation model of van Kampen, Rimes & Peacock (in preparation), a revised version of the model of van Kampen, Jimenez & Peacock (1999), is semi-numerical, in the sense that the merging history of galaxy haloes is taken directly from *N*-body simulations that include special techniques to prevent galaxy-scale haloes undergoing 'over-merging' owing to inadequate numerical resolution. When haloes merge, a criterion based on dynamical friction is used to decide how many galaxies exist in the newly merged halo. The most massive of those galaxies becomes the single central galaxy to which gas can cool, while the others become its satellites.

When a halo first forms, it is assumed to have an isothermal sphere density profile. A fraction $\Omega_{\mathrm{b}}/\Omega_{\mathrm{m}}$ of this is in the form of gas at the virial temperature, which can cool to form stars within a single galaxy at the centre of the halo. Application of the standard radiative cooling curve shows the rate at which this hot gas cools and is able to form stars. Energy output from supernovae reheats some of the cooled gas back to the hot phase. When haloes merge, all hot gas is stripped and ends up in the new halo. Thus each halo maintains an internal account of the amounts of gas being transferred between the two phases, and consumed by the formation of stars.

The model includes two modes of star formation: quiescent star formation in discs, and starbursts during major merger events. Having formed stars, in order to predict the appearance of the resulting galaxy it is necessary to assume an initial mass function (IMF), which is generally taken to be Salpeter's, and to have a spectral synthesis code, for which we use the spectral models of Bruzual & Charlot (1993). The evolution of the metals is followed, because the cooling of the hot gas depends on metal content, and a stellar population of high metallicity will be much redder than a low-metallicity one of the same age. It is taken as established that the population of brown dwarfs makes a negligible contribution to the total stellar mass density, and the model does not allow an adjustable mass-to-light ratio for the stellar population. The cosmological model adopted is $\Omega_{\mathrm{m}} = 0.3$, $\Omega_{\Lambda} = 0.7$, $h = 0.7$, $\sigma_8 = 0.93$, $\Omega_{\mathrm{b}} = 0.02\,h^{-2}$. The 850-μm flux is assumed to be proportional to the star formation rate [with 8 mJy corresponding to 1000 $\mathrm{M}_{\odot}\ \mathrm{yr}^{-1}$, as found by Scott et al. (2002)], with a random term of the order of 50 per cent added or subtracted to mimic the uncertainty in dust temperature, grain sizes and other properties that are not yet included in the modelling.

The model used in this paper has a mixture of bursting and quiescent star formation, with most of the recent star formation occurring in discs, following the Schmidt law with a threshold according to the Kennicutt criterion, and most of the high-redshift star formation resulting from merger-driven starbursts. The model is similar in philosophy to that of Hatton et al. (2003) and Baugh et al. (2005), but with many differences in the details and parameters adopted.

3.5 Model comparison

In this section we compare the models in order to get a qualitative description of where the differences lie.

What drives the flux in each of the models? In the simple merger model, an actual flux is not calculated; rather, a merger mass threshold is related to a flux threshold. The hydrodynamical and stable clustering models have some of the properties of the galaxies fixed, but not the flux, which is assigned statistically. The phenomenological model is the only one that generates a flux from the actual physical properties of the galaxies, taking into account the approximations and assumptions made.

What drives the clustering signal in each of the models? In the stable clustering model the *spatial* correlation function is fixed in comoving space, which means that the angular correlation function is built up along the light-cone in a way that depends on the selection of galaxies as a function of redshift. The simple merger, hydrodynamical and phenomenological models produce galaxies first, and build up the angular correlation function along the line of sight. In the simple merger a one-to-one correspondence between galaxies

and haloes is assumed, whereas the other two models have more complex relations between mass and light.

What determines the redshift distribution in each of the models? In the hydrodynamical model it is simply taken from Baugh et al. (1996), whereas for the other models it is actually an outcome of the models, although in the stable clustering model an exponential cut-off at $z \approx 6$ is applied. The main determining factor for the simple merger and phenomenological models for the redshift distribution is the merger rate as a function of redshift. For the simple merger model this is obvious, but for the phenomenological model this follows from the dominant contribution of merger-driven starbursts to the submillimetre flux.

Besides the models used in this paper, other models exist in the literature, which are similar in philosophy to those included here, but still produce different predictions for the submillimetre population. Models similar to the phenomenological model are those of Hatton et al. (2003) and Baugh et al. (2005), which differ mainly in the details of the physics implemented, and the choice of parameters.

4 MODEL PREDICTIONS COMPARED

In Fig. 1 we show a complete simulation of one of the two 850-μm data sets that will comprise SHADES, produced using the phenomenological model described in Section 3.4. A simple square geometry was chosen, although the actual survey geometry of each of the two SHADES fields could be of a somewhat different shape. All sources with fluxes larger than 8 mJy are shown, where the symbol size is proportional to the logarithm of the flux, i.e. a sub-millimetre magnitude.

We now consider simulated SHADES data sets as predicted from the various alternative models of the submillimetre source population presented in Section 3. We do not consider the effects of noise and sidelobes, but we do take into account the effects of the 15-arcsec SCUBA beam by merging into single submillimetre sources anything closer together than 7 arcsec. This reflects the resolution expected for the final source extraction from the SHADES images. After the removal of close pairs we assume the model sub-millimetre sources to reflect the final SHADES source list.

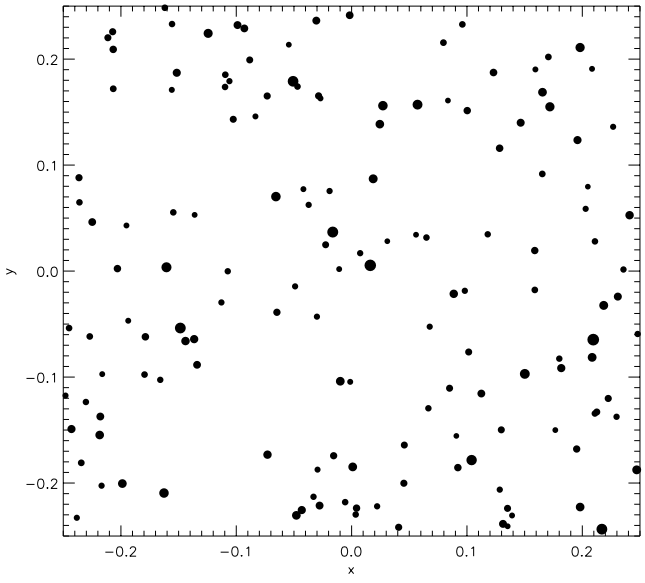

Figure 1. A simulated distribution of 8-mJy sources for the phenomenological model described in Section 3.4, with no redshift selection. The size of the field is equal to each of the two fields that will make up SHADES: 0.25 deg^2. The diameter of each dot is proportional to the submillimetre magnitude of the source that it represents.

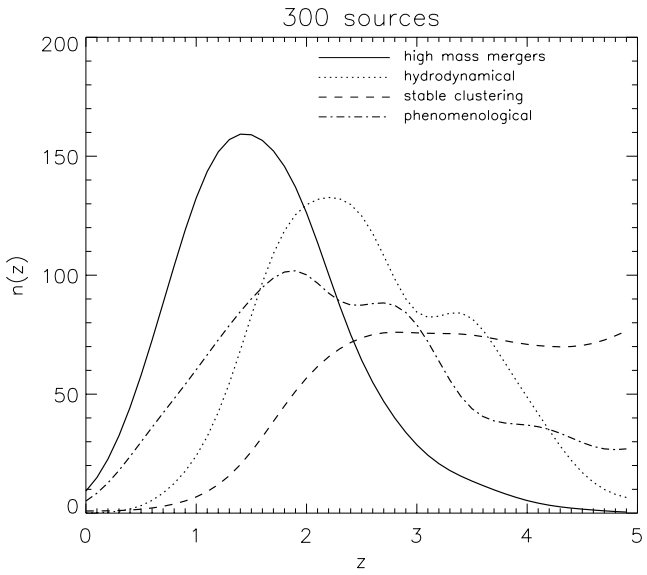

Figure 2. Redshift distributions for all models with 300 sources, averaged over 25 mock SHADES data sets for each model. The model distributions were convolved with a Gaussian of radius 0.4, roughly mimicking the statistical uncertainty in the photometric redshift errors.

The final survey is expected to contain around 300 sources, i.e. 150 sources per field. We produce, for each model, 50 realizations of individual fields, i.e. 25 mock SHADES data sets of 300 sources each.

4.1 Predictions for the redshift distribution

For the four models, we show in Fig. 2 the redshift distributions expected after smoothing with a Gaussian filter of radius 0.4, which reflects, very crudely, the resolution achievable with photometric redshifts. The distributions are obtained by averaging over all realizations for each model, and are normalized to the total source count of 300.

Even after relatively heavy smoothing, we see that the redshift distributions are rather different, and are clearly distinguishable from each other. This means that even crude but complete redshift information will be of enormous benefit in differentiating between and constraining models. Obviously, obtaining more accurate redshifts should help to tune the models that survive this first test even further.

For the purpose of comparing clustering properties between the models, note from Fig. 2 that the redshift range $2 < z < 3$ is the only range where all models have a reasonable number of sources to attempt a correlation function analysis. The differences in the distribution stem from the different assumptions for each of the models: the simple merger model assumes that only high-mass mergers can form SCUBA sources, which, in a hierarchical structure formation scenario, necessarily places them at lower redshifts as compared with the other models, while the simple, unbiased galaxy formation prescription of the stable clustering model places SCUBA sources at relatively high redshifts.

Current spectroscopic redshift measurements for submillimetre-selected galaxies are incomplete and only available for small

samples, so any redshift distribution derived from such measurements is tentative. Chapman et al. (2005) claim that a Gaussian distribution with $\bar{z} = 2.4$ and $\sigma_z = 0.65$ fits their available data well, but the incompleteness of their data set imposed through their radio selection hinders a comparison with models.

4.2 Clustering measures

The estimated redshifts have a predicted accuracy of $\delta z \sim \pm 0.4$, which means that we cannot directly measure the three-dimensional spatial correlation function $\xi(r)$. However, we do not have to restrict ourselves to measuring angular clustering, as photometric redshifts can be used to boost the angular clustering signal-to-noise ratio by splitting the sample in redshift bins, or by considering only pairs of galaxies that lie at similar redshifts. Even so, the measured correlation function will be noisy, so we use integrals of this function, as considered in the early days of optical galaxy surveys when total source counts were much lower than today (e.g. Davis & Peebles 1983).

4.2.1 Estimating the angular correlation function

The method for modelling the clustering of sources proceeds as follows. From the data, the Landy & Szalay (1993) estimator $1 + w_{\rm LS} = 1 + (DD - 2DR + RR)/RR$ is calculated, where *DD*, *DR* and *RR* are the (normalized) galaxy–galaxy, galaxy–random and random–random pair counts at separation θ, calculated from the galaxy sample and a large random catalogue containing 10 000 points that Poisson samples the survey region. This estimator is then fitted by its expected value

$$1 + \langle w_{\rm LS} \rangle = [1 + w(\theta)]/(1 + w_\Omega), \tag{1}$$

where w_Ω is the integral of the model two-point correlation function over the sampling geometry:

$$w_\Omega = \int_\Omega G_{\rm p}(\theta) w(\theta)\, {\rm d}\Omega. \tag{2}$$

The function $G_{\rm p}(\theta)$ is the probability density function of finding two randomly placed points in the survey at a distance θ. This 'integral constraint' corrects for the effect of not knowing the true density of objects (Groth & Peebles 1977; Landy & Szalay 1993) and stops the recovered correlation function being biased to low values compared with the true function. Note that equation (1) implies that the true correlation function is biased low by a *factor* of $1 + w_\Omega$, whereas often this is approximated by $w(\theta) = w_{\rm LS} - w_\Omega$, i.e. ignoring the term $w_{\rm LS} w_\Omega$.

We also introduce an alternative to the standard angular correlation function that takes redshift information into account in an unorthodox way. In the counting of *DD* pairs, we just consider those pairs that have a redshift separation of at most 0.4, whereas the *DR* and *RR* counts are still obtained for all galaxy pairs. This is equivalent to removing distant pairs that are expected to be unclustered from an analysis of the angular correlation function of all of the objects in the survey. It is clear that this approach must increase the signal-to-noise ratio of the recovered correlation function.

4.2.2 The sky-averaged angular correlation function

For the relatively small number of sources being detected in SHADES, we measure an integral of the Landy & Szalay estimator. Such an approach has previously been used to analyse clustering within early galaxy redshift surveys (e.g. the CfA survey: Davis & Peebles 1983). The statistic that was often obtained in these analyses was the integrated quantity J_3, defined as

$$J_3(r) \equiv \int_0^r \xi(y) y^2\, {\rm d}y. \tag{3}$$

The dimensionless analogue of J_3 is called the volume-averaged correlation function:

$$\bar{\xi}(r) \equiv \frac{3}{r^3} \int_0^r \xi(y) y^2\, {\rm d}y = 3\frac{J_3(r)}{r^3}. \tag{4}$$

This measures the fluctuation power up to the scale r, and is therefore a useful measure for a survey that is limited in object numbers. For reference, $\bar{\xi}(10\ h^{-1}\,{\rm Mpc}) = 0.83$ was found for the optical CfA survey (Davis & Peebles 1983; no error given).

As we cannot measure the spatial clustering function $\xi(r)$, as the redshift determinations are very uncertain, we use the two-dimensional version of $\bar{\xi}$, the sky-averaged angular correlation function

$$\bar{w}(\theta) \equiv \frac{2}{\theta^2} \int_0^\theta w(\phi)\phi\, {\rm d}\phi, \tag{5}$$

where $w(\theta)$ is the angular correlation function, which is the projection of $\xi(r)$ along the line of sight. Our estimator of this statistic was calculated by numerically integrating the angular correlation function [calculated using the Landy & Szalay (1993) estimator], in the form of logarithmically binned estimates w_i, up to the angle θ_i using equation (5):

$$\bar{w}_i = \frac{2}{\theta_i^2} \sum_{j \leqslant i} w_j \theta_j^2 \Delta, \tag{6}$$

where Δ is the logarithmic bin-size. The errors on $\bar{w}_i$ are obtained by propagating the errors on w_i through this summation. This estimate for the true sky-averaged angular correlation function is also biased, and has its own integral constraint, similar to the one for $w(\theta)$. For a power-law correlation function $w_{\rm pl}(\theta) = (\theta/A)^{-\delta}$,

$$\bar{w}_{\rm pl}(\theta) = \frac{2}{2 - \delta} \left(\frac{\theta}{A}\right)^{-\delta}. \tag{7}$$

Thus the sky-averaged correlation function is also a power law, with the same slope and a different amplitude (except for $\delta = 2$), and the integral constraint $\bar{w}_\Omega$ is scaled by the same factor $2/(2 - \delta)$ with respect to w_Ω (see equation 2). We can therefore fit power-law models to either $\bar{w}(\theta)$ or $w(\theta)$, and constrain the same parameters.

4.2.3 Fitting to the model correlation functions

Traditionally, χ^2 minimization is used to fit a power-law function to the estimated correlation function. This type of minimization is, strictly speaking, only valid for binned data that are uncorrelated and have Gaussian errors. Our estimates of $w(\theta)$ and $\bar{w}(\theta)$ are correlated for different maximum separations, and have errors that are non-Gaussian [it is easy to see that, in the absence of clustering, the errors are strictly Poisson, as shown by Landy & Szalay (1993) for the estimator of $w(\theta)$]. In order to constrain models of the correlation function using our binned estimates, we should therefore perform a full likelihood calculation taking into account the potentially complex shape of the likelihood. Both $w(\theta)$ and $\bar{w}(\theta)$ obviously contain the same information, and should therefore result in the same likelihood surface for a given model.

The reason for fitting $\bar{w}(\theta)$ rather than $w(\theta)$ lies in the approximations that are made in estimating the likelihood. For the sky-averaged

angular correlation function, the bins are dependent on more pairs of galaxies than the corresponding direct estimator of the correlation function, so that the sky-averaged statistic will have a distribution closer to a Gaussian form. Switching from a direct estimate to a sky-averaged estimate increases correlations between data points. However, this can be taken into account properly in the fitting procedure by calculating the full covariance matrix for the binned data, which is then diagonalized by a unitary transformation to produce an alternative χ^2-statistic (e.g. Fisher et al. 1994). This statistic is subsequently employed to fit models to the data.

A single-parameter fit to the correlation function is often adopted assuming that $w(\theta) = (\theta/A)^{-0.8}$ (e.g. Roche et al. 1993; Daddi et al. 2000). It is not at all clear what the slope for high-redshift submillimetre sources is going to be, but it is easier to fit a one-parameter function for a small number of galaxies. In the following we consider both a one-parameter model with constrained power-law slope $\delta = 0.8$, and a two-parameter fit for the generic power-law $w(\theta) = (\theta/A)^{-\delta}$ to both $w(\theta)$ and $\bar{w}(\theta)$. We employ non-linear χ^2-fitting for both functions, using the Levenberg–Marquardt method (Press et al. 1988), which allows us easily to take into account the multiplicative integral constraint. For each fit the χ^2 probability Q is calculated using the incomplete gamma-function, and any fits with $Q < 0.1$ are discarded.

4.2.4 *Examples of mock angular correlation functions*

In Figs 3, 4 and 5 we show examples of $w(\theta)$ and $\bar{w}(\theta)$ for realizations of all models, for the case where we have no redshift information (Fig. 3) and for the case where we have (Figs 4 and 5). The *same* realizations have been used for all three figures. Circles show the angular correlation function $w(\theta)$, where open symbols indicate negative values, and error bars indicate Poisson errors. The best-fitting power law is shown as a solid line, *if* a fit was possible. If not, the line is simply omitted. The sky-averaged angular correlation function $\bar{w}(\theta)$ is shown using asterisks, and the best-fitting power law is shown as a dashed line (again, if a fit was successful).

Before continuing, please note that these examples are by no means meant to be representative – they are merely shown to demonstrate the difference between using $w(\theta)$ and $\bar{w}(\theta)$ for fitting, and to show the effect of the additional redshift information. The examples should *not* be used to compare the differences in clustering strength between the models; this will be covered in the next section.

First focusing on Fig. 3, we see that the correlation functions for the complete line of sight are noisy, and for one of the models a fit to $w(\theta)$ fails completely. The figure demonstrates the use of sky-averaging, as $\bar{w}(\theta)$, plotted using stars, is better behaved. This is perhaps best illustrated in the first panel, but also in the third panel, where it has the direct consequence that a fit to $\bar{w}(\theta)$ is possible where a fit to $w(\theta)$ failed (third panel of Fig. 3). In general, for most realizations a simple χ^2 fit to $\bar{w}(\theta)$ turns out to be easier than a fit to the angular correlation function itself. This is very helpful for our purpose of comparing models, as it is important that fitting is possible for a large number of realizations, so the fitting procedure needs to be largely automatic.

A stronger clustering signal is expected for sources selected in a redshift range, as the signal is less polluted by uncorrelated sources at very different redshifts. Indeed, Fig. 4 shows that for all realizations the selection of sources in the redshift interval $2 < z < 3$ boosts

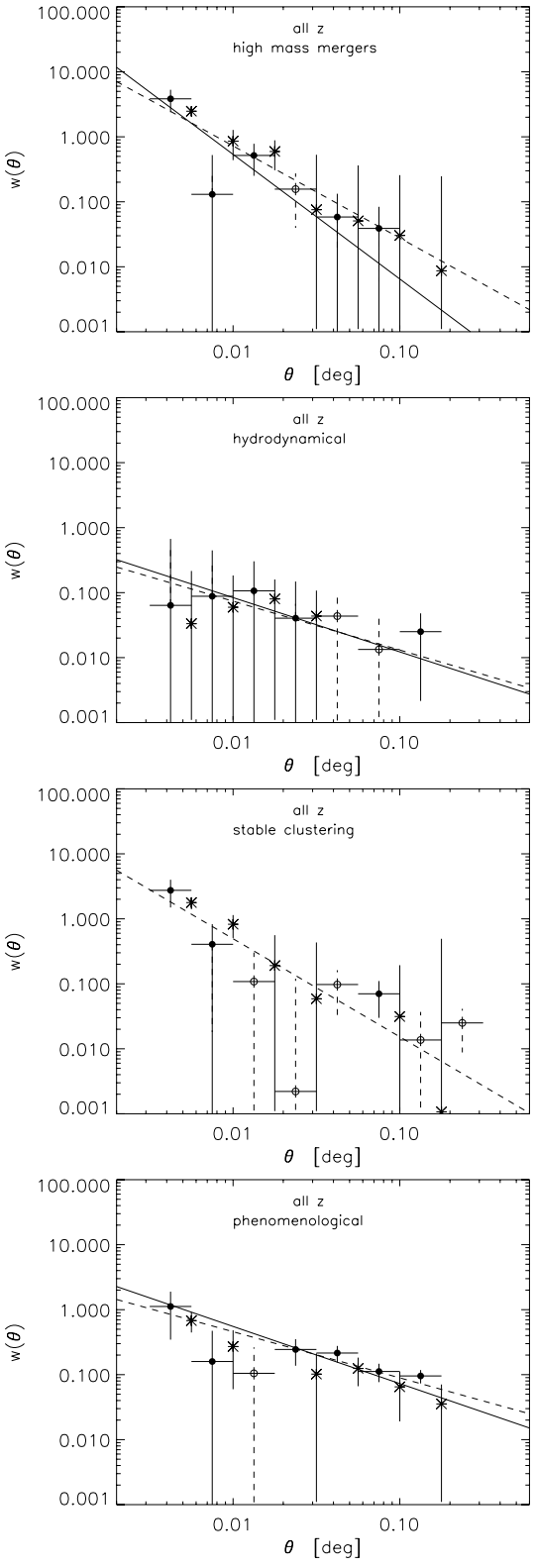

Figure 3. Bias-corrected angular correlation function $w(\theta)$ (circles, open denoting negative) and its sky-averaged counterpart $\bar{w}(\theta)$ (asterisks, negative values not plotted) for a single realization of each model, with best-fitting power-law functions overplotted for both, and no redshift selection. See text for full details.

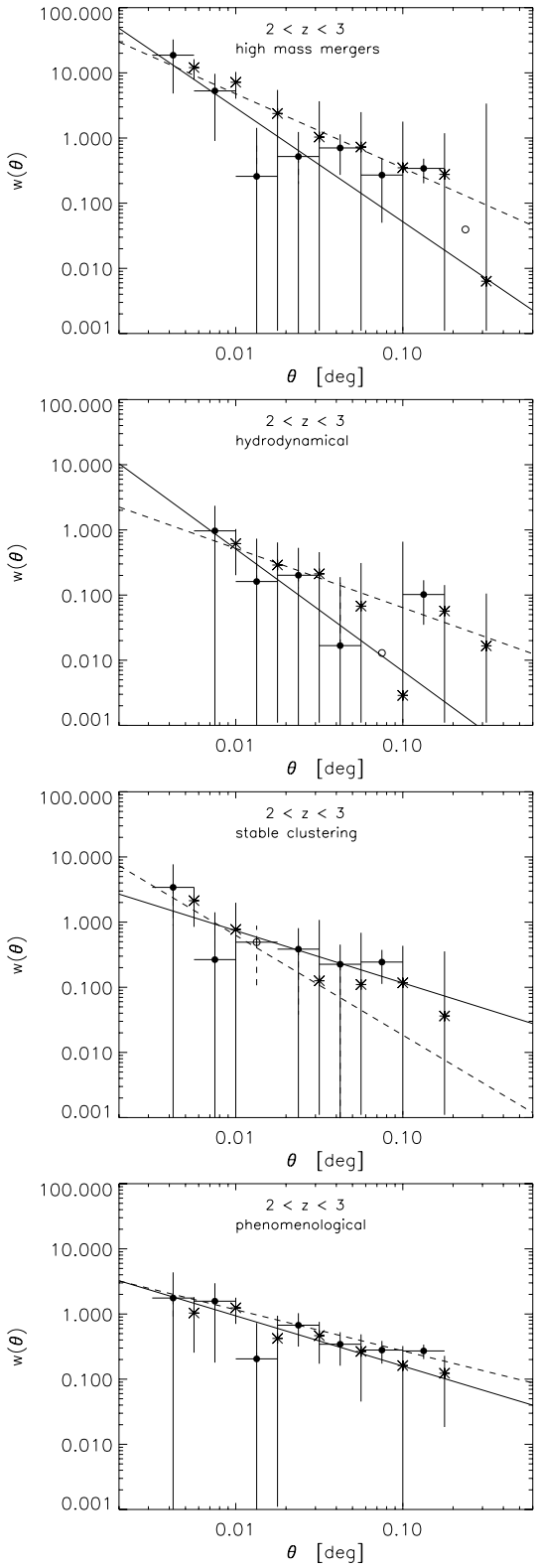

Figure 4. As Fig. 3, but with sources selected to be in the redshift range $2 < z < 3$, which is the range where all models have a reasonable number of sources (see Fig. 2). All realizations are exactly the same as in Fig. 3, in order to demonstrate what redshift availability can achieve.

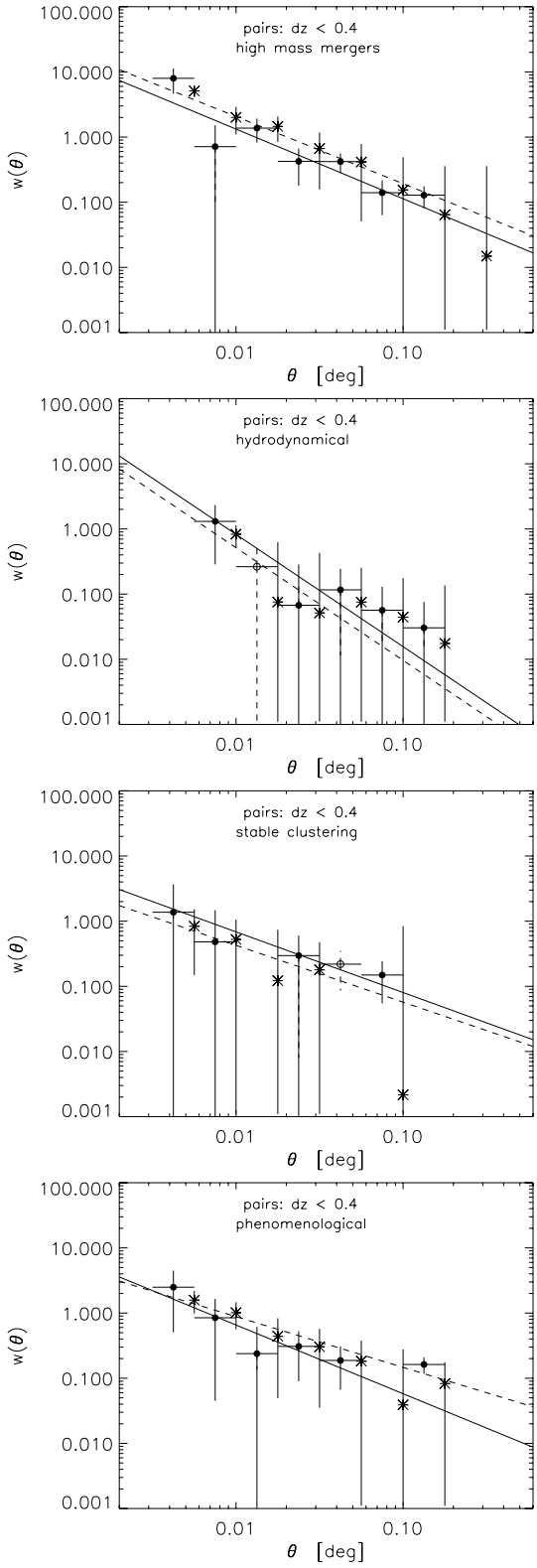

Figure 5. As Fig. 3, but with sources selected to be in redshift pairs with $\delta z < 0.4$. All realizations are exactly the same as in Figs 3 and 4.

the clustering signal. Even though the errors are larger because of the small number of sources, the stronger signal means that a fit is possible in all cases shown, even for $w(\theta)$ itself, although the sky-averaged correlation function is to be preferred

nevertheless. However, a fraction of realizations still produce unacceptable fits.

An alternative to redshift intervals is to count only galaxies that are paired in redshift space, e.g. have $|z_i - z_j| < 0.4$, as described in Section 4.2.1 The result for the same realizations as used for Figs 3 and 4 is shown in Fig. 5, and again much better results are obtained as compared with the estimates without any redshift information (Fig. 3). The binned data look cleaner than those for the redshift intervals (Fig. 4), which is due to more galaxies being used in the *DD* counts. The fitting is therefore somewhat more reliable, as demonstrated by the small difference between fits to $w(\theta)$ and $\bar{w}(\theta)$.

4.2.5 Distribution over fitting parameters

So far we have considered single realizations for each model, which should really be treated as *examples* of how the final SHADES data set might appear. In order to be able to make a quantitative comparison possible between the actual final SHADES data set and all models, we need to find the probability distribution over the fitting parameters for each model given the survey constraints (area, flux limit, etc.).

We therefore produced 50 realizations for each model, and fitted a power-law correlation function to all of these. The resulting amplitudes $A_{\rm sky}$ and slopes $\delta_{\rm sky}$ of the sky-averaged correlation function $\bar{w}(\theta)$ are shown as scatter plots in Fig. 6, which shows the case where no redshift information is available, and in Figs 7 and 8, where we are able to split up the sample in redshift intervals (three of these are shown), or select pairs of galaxies in redshift space.

In the case of no redshift information, a significant fraction of the mock fields do not produce a correlation function that can be fitted by a power law, and the number of estimates in Fig. 6 is therefore less than 50 for each model. However, for each model still more than half of the realization allow a good fit so, crudely speaking, one would expect that at least one of the two observed SHADES fields should produce a good fit. All models spread out over a relatively large region of parameter space, and seem to overlap with each other for most of that region. This merely reflects the fact that whatever intrinsic correlation exists in the underlying submillimetre population is weakened by projection, which produces this large spread in fitting parameters and the significant overlap between the models. Only the high-mass merger model shows a larger clustering

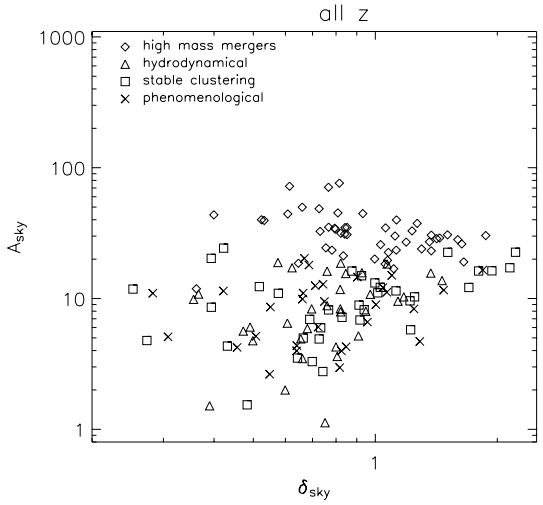

Figure 6. Scatter plot for the two fitting parameters of the sky-averaged angular correlation function, $A_{\rm sky}$ and $\delta_{\rm sky}$, for 50 fields of 150 sources. Only fits of sufficient quality, i.e. those with a χ^2 probability larger than 0.1, are included (see text for details).

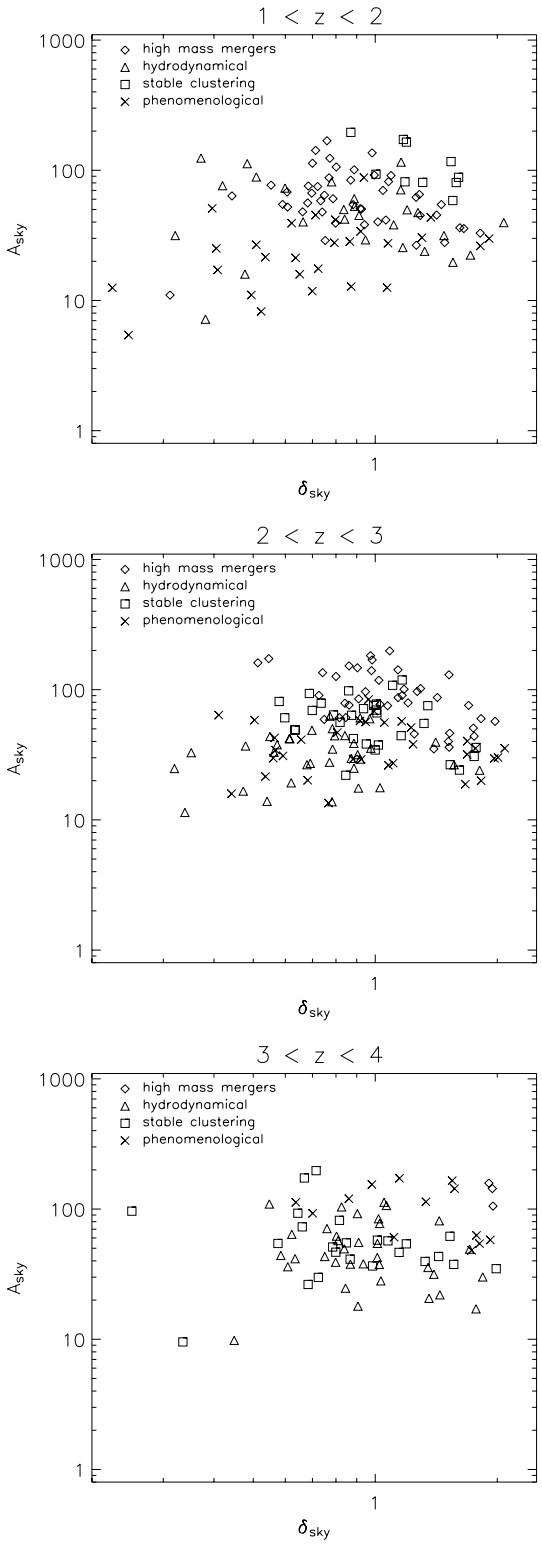

Figure 7. As Fig. 6, but for three different redshift bins. Again only parameters from 'good' fits are shown, i.e. those with a χ^2 probability larger than 0.1.

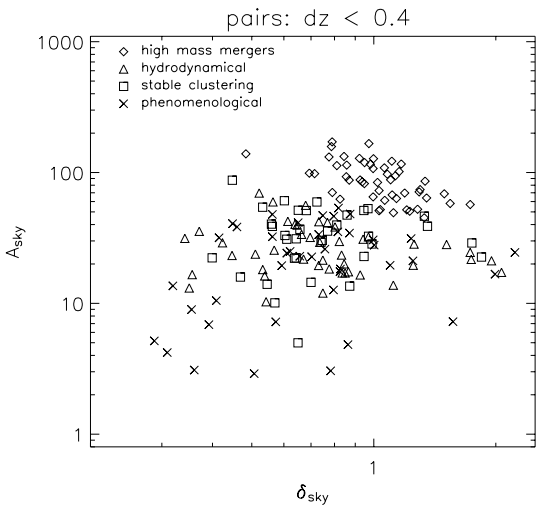

Figure 8. As Fig. 6, but for close redshift pairs with $\delta z < 0.4$. Again only parameters from 'good' fits are shown, i.e. those fits that have a χ^2 probability larger than 0.1.

amplitude overall, which also results in a larger fraction of the realizations producing good fits, and a smaller spread in the slope δ.

Let us now use the redshift information to split up the mock samples into redshift intervals, which should show stronger clustering. Three intervals are shown in Fig. 7, where the $2 < z < 3$ case is the most relevant one as it has the most sources for all models. The other two intervals both have at least one model where the number of sources is significantly lacking, which means that only the remaining models can reasonably be compared. The first thing to notice in all three panels of Fig. 7 is that the clouds of fitted parameter pairs start to separate out somewhat, reducing the overlap between the models.

Various interesting effects can be seen for the different models. The stable clustering model (open squares) shows fairly strong clustering in the $1 < z < 2$ redshift interval, but, because of the low number of sources in this redshift range (see Fig. 2), few of the realizations actually produce a good fit. The high-mass merger model (open diamonds) shows relatively strong clustering for $2 < z < 3$, but of course lacks numbers in the highest redshift interval. The phenomenological model (crosses) shows strongest clustering in $3 < z < 4$ interval, but this is also somewhat troubled by low source counts. It also shows the weakest clustering for the lowest redshift interval. The clustering strength of the hydrodynamical model (open triangles) is virtually independent of redshift.

In Fig. 8 we show the results for the close pairs, i.e. galaxies with $|\delta z| < 0.4$. The points scatter in a similar fashion to the $2 < z < 3$ interval, with some differences, and the separating of model point clouds is comparable between models, except for the simple merger model, which can quite clearly be distinguished. This diagram should thus be a good test for high-mass merging versus the other models considered in this paper.

If we concentrate on the $2 < z < 3$ interval, which allows the cleanest comparison between the four models considered here, and on the close pairs, we do see that the clouds of points are overlapping significantly, but the distributions are elongated somewhat along the δ-axis, and have different mean clustering amplitudes A. Also, the mean of the distribution is near $\delta = 0.8$, observationally found for a range of galaxy types.

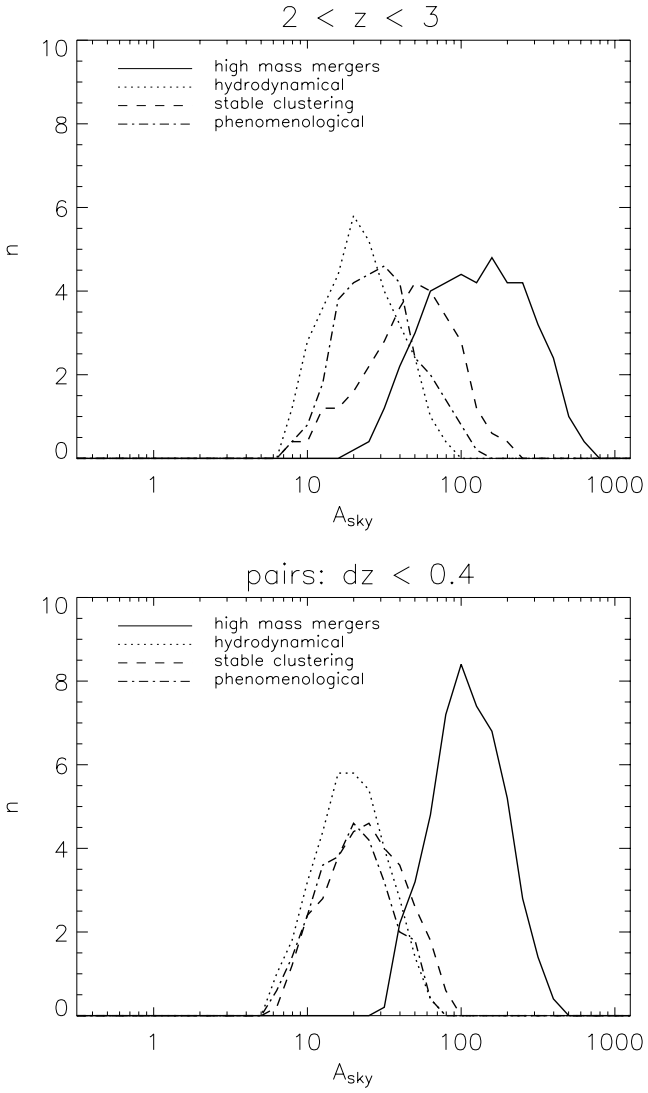

Figure 9. Distribution of clustering amplitude $A_{\rm sky}$ over 50 realizations for each model, for the redshift bin $2 < z < 3$, where each model has a sufficient number of sources available (top panel), and for close redshift pairs with $|\delta z| < 0.4$ (bottom panel).

This leads us finally to consider the traditional one-parameter fit to the data, assuming $\delta = 0.8$. This produces a single clustering amplitude $A_{\rm sky}$ for each mock field, and a distribution over $A_{\rm sky}$ for each model. These distributions are plotted in Fig. 9, for the redshift interval $2 < z < 3$ (top panel) and for the close pairs. Interestingly, for this redshift interval, all distributions are different, and although there is significant overlap the final SHADES data set will distinguish between the models, especially in combination with the different redshift distributions (see Fig. 2). For the close pairs, the result of Fig. 8 is made more apparent, in that the high-mass merger model is clearly different from the rest of the models, which show almost identical distributions.

Another measure of clustering is the sky-averaged angular correlation function as it was originally intended: just as a measurement. Therefore we also plot, in Fig. 10, the distribution over a particular $\bar{w}(\theta_i)$, which we choose to be $\bar{w}(1$ arcmin), i.e. the sky-averaged correlation function within 1 arcmin, for the same redshift-selected data as used for Fig. 9. The result is similar, although the distributions overlap more than those for $A_{\rm sky}$ (as seen in Fig. 9). However,

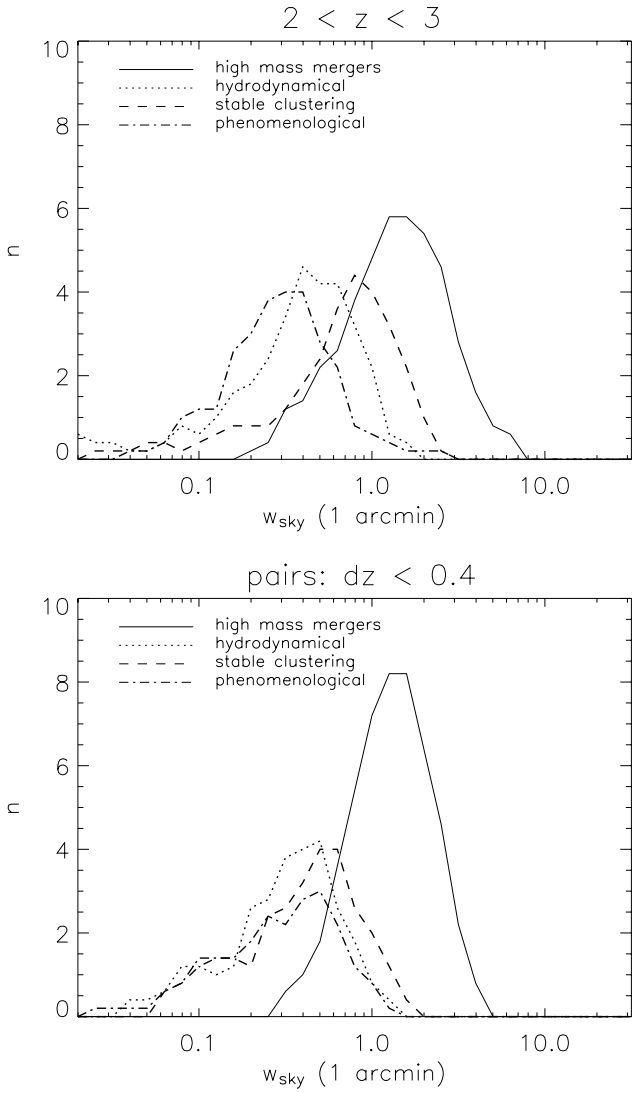

Figure 10. Distribution of the angular correlation measure $\bar{w}(1\ \mathrm{arcmin})$ over 50 realizations for each model, again for the redshift bin $2 < z < 3$ (top panel), and the close redshift pairs with $\delta z < 0.4$ (bottom panel).

the major advantage with the measure $\bar{w}(1\ \mathrm{arcmin})$, or one at a different angle, is that we do not need to assume a model for the form of the correlation function.

5 CONCLUSIONS

One of the primary science drivers for the SHADES project is to place strong constraints on galaxy formation models by observations of luminous submillimetre galaxies in the high-redshift Universe. In order to achieve this, it is worth considering how models are best constrained by the data, and examine the range of possible predictions. With this aim in mind, we have presented four different models of the submillimetre galaxy population, selected to be widely varying in concept, without worrying about every aspect of each model. We avoid any strong assumption about the nature or redshift distribution of the submillimetre population. Given the uncertainties in what is known about the submillimetre population, we want to keep open a range of models, even those that can be questioned in some of their aspects.

For each model, the redshift distribution and clustering properties of the submillimetre population have been predicted for SHADES, and 50 realizations have been produced, each comprising around 150 sources. Thus 25 mock SHADES data sets have been produced for each model. These simulated SHADES catalogues have been used to investigate the *ability* of the clustering statistics of the final data set to constrain the various models collected here. Direct and sky-averaged estimators of the correlation function have been considered and their relative merits discussed. We have argued that power-law fits are best performed on the sky-averaged angular correlation function, and that a relatively good fit is possible in most cases for this measure. The full covariance matrix has been calculated and diagonalized, resulting in a χ^2 statistic that has been used to fit the power-law model to the data using the Levenberg–Marquardt method as implemented by Press et al. (1988).

All models predict sufficiently strong clustering, so that we expect to detect clustering within the SCUBA population when SHADES is complete. Although cosmic variance remains a concern, we can certainly quantify the probability of a given model to produce the observed data set, and this is expected to reject some of the models included in this paper. However, the aim of this paper is not to constrain the models; this will be done when the full SHADES data set is available. In fact, we simply have observed basic trends between models by reducing each measured correlation function to two power-law parameters: its slope and amplitude respectively.

The observed redshift distribution will provide a complementary strong test of models, even with relatively coarse photometric redshift information. In fact, the combination of clustering and redshift data offers the best discriminator between the different models that we have considered: models with similar redshift distributions have different clustering strengths, while models with similar clustering properties have different redshift distributions.

Recently, Blain et al. (2004) considered the clustering of submillimetre sources in a number of relatively small fields for data with follow-up spectroscopic redshifts. An approach was adopted that selected galaxy pairs based only on radial positions, and did not use any angular information. Pairs of submillimetre sources with $\Delta v = 1200\ \mathrm{km\ s^{-1}}$, equivalent to separations of the order of 5 Mpc (comoving) at $z = 2.5$, were counted, and compared with the integrated model $\xi(r)$. For a larger survey where there is significant angular information, this method is not optimal, so for SHADES we prefer fully to exploit the angular information, with the restricted photometric redshift information as a subsample selection tool.

The Blain et al. (2004) method is equivalent to the standard method of calculating $\xi(r)$ by pair counting (using $DD/RR - 1$), but performing this in a single bin with $r < R_{\mathrm{max}}$. While angular clustering measurements ignore radial information, this 'radial' method instead ignores angular information in a similar way. The method should therefore include the equivalent of an integral constraint. Also, the quoted errors are derived from Poisson errors (on the pair counts), whereas the errors are expected to be larger than Poissonian for this estimator (Landy & Szalay 1993). Further limitations of the method employed by Blain et al. (2004) are discussed by Adelberger (2005). Given these issues, it is probably too early to use the Blain et al. clustering results to distinguish between models.

The primary conclusion from our analysis is that, with the area coverage ($0.5\ \mathrm{deg}^2$) and the expected number of sources (200–400), and particularly with the expected photometric redshift information ($\Delta z \sim \pm 0.3$), SHADES is capable of distinguishing between widely varying scenarios for the production of the bright submillimetre population.

ACKNOWLEDGMENTS

This research was supported in part by the Austrian Science Foundation FWF under grant P15868, by two PPARC rolling grants, and by the HPC Europa programme funded through the European Commission contract number RII3-CT-2003-506079. We thank Duncan Farrah and Tom Babbedge for many useful comments and suggestions. DHH and IA are partly supported by CONACyT grants 39953-F and 39548-F.

REFERENCES

Adelberger K. L., 2005, ApJ, in press (astro-ph/0412397)
Adelberger K. L., Steidel C. C., 2000, ApJ, 544, 218
Almaini O. et al., 2003, MNRAS, 338, 308
Aretxaga I., Hughes D. H., Chapin E. L., Gaztañaga E., Dunlop J. S., Ivison R. J., 2003, MNRAS, 342, 759
Aretxaga I., Hughes D. H., Dunlop J. S., 2005, MNRAS, in press (doi:10.1111/j.1365-2966.2005.08733.x)
Barger A. J., Cowie L. L., Sanders D. B., 1999, ApJ, 518, L5
Baugh C. M., Cole S., Frenk C. S., 1996, MNRAS, 282, 27
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, ApJ, 611, 725
Bruzual A. G., Charlot S., 1993, ApJ, 405, 538
Chapman S. C., Blain A. W., Ivison R. J., Smail I. R., 2003, Nat, 422, 695
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, in press
Couchman H. M. P., Thomas P. A., Pearce F. R., 1995, ApJ, 452, 797
Daddi E., Cimatti A., Pozzetti L., Hoekstra H., Röttgering H. J. A., Renzini A., Zamorani G., Mannucci F., 2000, A&A, 361, 535
Davis M., Peebles P. J. E., 1983, ApJ, 267, 465
Devlin M., 2001, in Lowenthal J. D., Hughes D. H., eds, Proc. UMass/INAOE Conf., Deep Millimeter Surveys: Implications For Galaxy Formation And Evolution. World Scientific Publishing, Singapore, p. 59
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J., 1994, MNRAS, 266, 50
Gaztañaga E., Baugh C. M., 1998, MNRAS, 294, 229
Gaztañaga E., Hughes D., 2001, in Lowenthal J. D., Hughes D. H., eds, Proc. UMass/INAOE Conf., Deep Millimeter Surveys: Implications For Galaxy Formation And Evolution. World Scientific Publishing, Singapore, p. 131
Giavalisco M., Dickinson M., 2001, ApJ, 550, 177
Groth E. J., Peebles P. J. E., 1977, ApJ, 217, 385
Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, MNRAS, 343, 75
Hughes D. H., Gaztañaga E., 2000, in Favata F., Kaas A., Wilson A., eds, ESA SP-445, Proc. 33rd ESLAB Symp., Star Formation From The Small To The Large Scale. ESA, Noordwijk, p. 29
Hughes D. H. et al., 1998, Nat, 294, 341
Hughes D. H. et al., 2002, MNRAS, 335, 871
Ivison R. et al., 2002, MNRAS, 337, 1
Kaiser N.,1984, ApJ, 284, 9
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Mortier A. et al., 2005, MNRAS, submitted
Muanwong O., Thomas P. A., Kay S. T., Pearce F. R., 2002, MNRAS, 336, 527
Percival W. J., Miller L., Peacock J. A., 2000, MNRAS, 318, 273
Percival W. J., Scott D., Peacock J. A., Dunlop J. S., 2003, MNRAS, 338, 31
Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T., 1988, Numerical Recipes: The Art of Scientific Computing. Cambridge Univ. Press, Cambridge
Roche N., Shanks T., Metcalfe N., Fong R., 1993, MNRAS, 262, 456
Scott S. E. et al., 2002, MNRAS, 331, 817
Springel V., Yoshida N., White S. D. M., 2001, New Astron., 6, 79
Thacker R. J., Tittley E. R., Pearce F. R., Couchman H. M. P., Thomas P. A., 2000, MNRAS, 319, 619
van Kampen E., Jimenez R., Peacock J. A., 1999, MNRAS, 310, 43
Webb T. M. et al., 2003, ApJ, 587, 41

This paper has been typeset from a TEX/LATEX file prepared by the author.